

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
23 January 2003 (23.01.2003)

PCT

(10) International Publication Number  
**WO 03/006477 A1**

- (51) International Patent Classification<sup>7</sup>: C07H 21/02, 21/04, C12N 5/00, A01K 67/00 (74) Agent: FASSE, J., Peter; Fish & Richardson P.C., 225 Franklin Street, Boston, MA 02110-2804 (US).
- (21) International Application Number: PCT/US02/22010 (81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (22) International Filing Date: 12 July 2002 (12.07.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/305,185 12 July 2001 (12.07.2001) US (84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (71) Applicant (*for all designated States except US*): UNIVERSITY OF MASSACHUSETTS [US/US]; One Beacon Street, 26th Floor, Boston, MA 02108 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (*for US only*): ZAMORE, Phillip, D. [US/US]; 500 Green Street, Northborough, MA 01532 (US). HUTVAGNER, Gyorgy [HU/US]; 4 Browker Street #3, Worcester, MA 01604 (US). MCLACHLAN, Juanita [US/US]; 1199 Grafton Street #88F, Worcester, MA 01604 (US). MELLO, Craig, C. [US/US]; 19 Ryan Road, Shrewsbury, MA 01545 (US). GRISHOK, Alla [UA/US]; 263 Watertown Street, Newton, MA 02458 (US).
- Published:**  
— with international search report  
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: IN VIVO PRODUCTION OF SMALL INTERFERING RNAs THAT MEDIATE GENE SILENCING

(57) Abstract: The invention provides engineered RNA precursors that when expressed in a cell are processed by the cell to produce targeted small interfering RNAs (siRNAs) that selectively silence targeted genes (by cleaning specific mRNAs) using the cell's own RNA interference (RNAi) pathway. By introducing nucleic acid molecules that encode these engineered RNA precursors into cells *in vivo* with appropriate regulatory sequences, expression of the engineered RNA precursors can be selectively controlled both temporally and spatially, i.e., at particular times and/or in particular tissues, organs, or cells.

WO 03/006477 A1

## IN VIVO PRODUCTION OF SMALL INTERFERING RNAs THAT MEDIATE GENE SILENCING

### TECHNICAL FIELD

5 This invention relates to ribonucleic acid interference (RNAi), and more particularly to RNAi *in vivo*.

### BACKGROUND

RNAi is the sequence-specific, post-transcriptional silencing of a gene's expression by double-stranded RNA. RNAi is mediated by 21 to 25 nucleotide, double-stranded RNA molecules referred to as small interfering RNAs (siRNAs) that  
10 are derived by enzymatic cleavage of long, double-stranded RNA in cells. siRNAs can also be synthesized chemically or enzymatically outside of cells and then delivered to cells (e.g., by transfection) (see, e.g., Fire et al., 1998, "Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*," Nature, 391:806-11; Tuschl et al., 1999, "Targeted mRNA degradation by double-stranded RNA in vitro," Genes Dev., 13:3191-7; Zamore et al., 2000, "RNAi: double-  
15 stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals," Cell, 101:25-33.; Elbashir et al., 2001, "Duplexes of 21-nucleotide RNAs mediate RNA interference in mammalian cell culture," Nature, 411:494-498; and Elbashir et al., 2001, "RNA interference is mediated by 21- and 22-nucleotide RNAs,"  
20 Genes Dev., 15:188-200.

Double-stranded siRNAs mediate gene silencing by targeting for disruption or cleavage messenger RNAs (mRNAs) that contain the sequence of one strand of the siRNA. siRNAs introduced into mammalian cells by transfection mediate sequence-specific gene silencing, whereas long, double-stranded RNA induces sequence non-  
25 specific responses.

### SUMMARY

The invention is based on the discovery of new artificial, engineered RNA precursors, that when expressed in a cell, e.g., *in vivo*, are processed by the cell to produce targeted siRNAs that selectively silence target genes (by targeting specific

mRNAs for cleavage) using the cell's own RNAi pathway. By introducing nucleic acid molecules that encode these engineered RNA precursors into cells *in vivo* with appropriate regulatory sequences (e.g., a transgene in a vector such as a plasmid), expression of the engineered RNA precursors can be selectively controlled both  
5 temporally and spatially, i.e., at particular times and/or in particular tissues, organs, or cells.

In general, the invention features an isolated nucleic acid molecule including a regulatory sequence operably linked to a nucleic acid sequence that encodes an engineered ribonucleic acid (RNA) precursor, wherein the precursor includes: (i) a  
10 first stem portion comprising a sequence of at least 18 nucleotides that is complementary to a sequence of a messenger RNA (mRNA) of a target gene; (ii) a second stem portion comprising a sequence of at least 18 nucleotides that is sufficiently complementary to the first stem portion to hybridize with the first stem portion to form a duplex stem (e.g., a stem that can be processed by the enzyme  
15 Dicer); and (iii) a loop portion that connects the two stem portions. In another aspect, the invention features the engineered RNA itself. The RNA precursor targets a portion of the mRNA of the target gene, disrupts translation of the mRNA by cleaving the mRNA, and thereby prevents expression of the protein to be inhibited. The target genes can be, for example, human genes, e.g., mutant human genes, e.g., having a  
20 point mutation, or they can be viral or other genes.

In these molecules and precursors, the first stem portion can be fully complementary (i.e., completely complementary) to the mRNA sequence. In other embodiments, the stem portion can be complementary, i.e., the sequence can be substantially complementary (e.g., there can be no more than one or two mismatches  
25 over a stretch of 20 nucleotides). Similarly, the second stem portion can fully or substantially complementary to the first stem portion. The first stem portion can be located at a 5' or 3' end of the RNA precursor.

In these precursors, the loop portion can include at least 4, 7, or 11, or more nucleotides, and the sequence of the mRNA is located from 100 to 300 nucleotides 3'  
30 of the start of translation of the mRNA. The sequence of the mRNA can be located in a 5' untranslated region (UTR) or a 3' UTR of the mRNA. The first and second stem portions can each include about 18 to about 30 nucleotides, or about 22 to about 28

nucleotides. The first and second stem portions can each have the same number of nucleotides, or one of the first and second stem portions can have 1 to 4 more nucleotides than the other stem portion. These overhanging nucleotides can all be uracils.

5 In these nucleic acid molecules, the regulatory sequence can be a Pol III or Pol II promoter, and can be constitutive or inducible. In specific embodiments, the engineered RNA precursor can have the sequence set forth in SEQ ID NO:1, 2, 3, 4, 5, 8, or 9, and the nucleic acid molecule can have the sequence set forth in SEQ ID NO:10, 11, 17, 18, 20, or 21, or a complement thereof.

10 In other embodiments, the invention also features vectors, e.g., plasmids or viral (e.g., retroviral) vectors, that include the new nucleic acid molecules.

In another aspect, the invention includes host cells, e.g., mammalian cells, that contain the new nucleic acid molecules. The invention also includes transgenes that include the new nucleic acid molecules.

15 In another aspect of the invention, the invention features transgenic, non-human animals, one or more of whose cells include a transgene containing one or more of the new nucleic acid molecules, wherein the transgene is expressed in one or more cells of the transgenic animal resulting in the animal exhibiting ribonucleic acid interference (RNAi) of the target gene by the engineered RNA precursor. For  
20 example, the transgene can be expressed selectively in one or more cardiac cells, lymphocytes, liver cells, vascular endothelial cells, or spleen cells. In these animals, the regulatory sequence can be constitutive or inducible, or the regulatory sequence can be tissue specific. In some embodiments, the regulatory sequence can be a Pol III or Pol II promoter, and can be an exogenous sequence. These transgenic animals can  
25 be non-human primates or rodents, such as mice or rats, or other animals (e.g., other mammals, such as goats or cows; or birds) described herein.

The invention also includes cells derived from the new transgenic animals. For example, these cells can be a lymphocyte, a hematopoietic cell, a liver cell, a cardiac cell, a vascular endothelial cell, or a spleen cell.

30 In another aspect, the invention includes methods of inducing ribonucleic acid interference (RNAi) of a target gene in a cell, e.g., in an animal or in culture. The new methods include obtaining a transgenic animal comprising a transgene including

a nucleic acid molecule encoding an engineered RNA precursor and an inducible promoter; and inducing the cell to express the precursor to form a small interfering ribonucleic acid (siRNA) within the cell, thereby inducing RNAi of the target gene in the animal.

5           Alternatively, the methods include obtaining a host cell; culturing the cell; and enabling the cell to express the RNA precursor to form a small interfering ribonucleic acid (siRNA) within the cell, thereby inducing RNAi of the target gene in the cell.

A "transgene" is any nucleic acid molecule, which is inserted by artifice into a cell, and becomes part of the genome of the organism that develops from the cell.

10          Such a transgene may include a gene that is partly or entirely heterologous (i.e., foreign) to the transgenic organism, or may represent a gene homologous to an endogenous gene of the organism. The term "transgene" also means a nucleic acid molecule that includes one or more selected nucleic acid sequences, e.g., DNAs, that encode one or more engineered RNA precursors, to be expressed in a transgenic  
15          organism, e.g., animal, which is partly or entirely heterologous, i.e., foreign, to the transgenic animal, or homologous to an endogenous gene of the transgenic animal, but which is designed to be inserted into the animal's genome at a location which differs from that of the natural gene. A transgene includes one or more promoters and any other DNA, such as introns, necessary for expression of the selected nucleic acid  
20          sequence, all operably linked to the selected sequence, and may include an enhancer sequence.

A "transformed cell" is a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a nucleic acid molecule or transgene encoding an engineered RNA precursor.

25          As used herein, the term "operably linked" means that a selected nucleic acid sequence, e.g., encoding an engineered RNA precursor, is in proximity with a promoter, e.g., a tissue-specific promoter, to allow the promoter to regulate expression of the selected nucleic acid sequence. In addition, the promoter is located upstream of the selected nucleic acid sequence in terms of the direction of transcription and  
30          translation.

By "promoter" is meant a nucleic acid sequence that is sufficient to direct transcription. A tissue-specific promoter affects expression of the selected nucleic

acid sequence in specific cells, e.g., hematopoietic cells, or cells of a specific tissue within an animal, e.g., cardiac, muscle, or vascular endothelium. The term also covers so-called "leaky" promoters, which regulate expression of a selected nucleic acid sequence primarily in one tissue, but cause expression in other tissues as well.

5 Such promoters also may include additional DNA sequences that are necessary for expression, such as introns and enhancer sequences.

By "transgenic" is meant any cell that includes a nucleic acid, e.g., DNA sequence, that is inserted by artifice into a cell and becomes part of the genome of an organism that develops from that cell. A "transgenic animal" means an animal that  
10 includes a transgene that is inserted into an embryonal cell and becomes a part of the genome of the animal which develops from that cell, or an offspring of such an animal. In the transgenic animals described herein, the transgene causes specific tissue cells to express an engineered RNA precursor. Any animal that can be produced by transgenic technology is included in the invention, although mammals  
15 are preferred. Preferred mammals include non-human primates, sheep, goats, horses, cattle, pigs, rabbits, and rodents such as guinea pigs, hamsters, rats, gerbils, and, preferably, mice.

An "isolated nucleic acid molecule or sequence" is a nucleic acid molecule or sequence that is not immediately contiguous with both of the coding sequences with  
20 which it is immediately contiguous (one on the 5' end and one on the 3' end) in the naturally occurring genome of the organism from which it is derived. The term therefore includes, for example, a recombinant DNA or RNA that is incorporated into a vector; into an autonomously replicating plasmid or virus; or into the genomic DNA of a prokaryote or eukaryote, or which exists as a separate molecule (e.g., a cDNA or  
25 a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other sequences. It also includes a recombinant DNA that is part of a hybrid gene encoding an additional polypeptide sequence.

A "target gene" is a gene whose expression is to be selectively inhibited or "silenced." This silencing is achieved by cleaving the mRNA of the target gene by an  
30 siRNA that is created from an engineered RNA precursor by a cell's RNAi system. One portion or segment of a duplex stem of the RNA precursor is an anti-sense strand

that is complementary, e.g., fully complementary, to a section of about 18 to about 40 or more nucleotides of the mRNA of the target gene.

The term “engineered,” as in an engineered RNA precursor, or an engineered nucleic acid molecule, indicates that the precursor or molecule is not found in nature, in that all or a portion of the nucleic acid sequence of the precursor or molecule is created or selected by man. Once created or selected, the sequence can be replicated, translated, transcribed, or otherwise processed by mechanisms within a cell. Thus, an RNA precursor produced within a cell from a transgene that includes an engineered nucleic acid molecule is an engineered RNA precursor.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

The invention provides several advantages. For example, the invention improves on and overcomes a significant deficiency in the prior art. Prior methods for inducing RNAi in mammalian cells using siRNAs were restricted to cell cultures. The new methods extend RNAi to whole animals, e.g., mammals, and thus allow RNAi to be targeted to specific cell types, organs, or tissues, and/or to specific developmental stages.

In addition, this technology simplifies and lowers the cost of siRNA construction, because DNA molecules are relatively inexpensive to make. Thus, large populations of plasmids or other vectors can be prepared, each containing a nucleic acid molecule that encodes an engineered RNA precursor that targets a particular gene, can be easily prepared, e.g., in an array format. In addition, the new nucleic acid molecules can be introduced into a variety of cells, which can be cultured *in vitro* using known techniques. Furthermore, the new methods enable the long-term, e.g.,

permanent, reduction of targeted gene expression in cell lines, because siRNAs are transient, but a transgenic hairpin provides a long-lasting supply of siRNAs.

The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

### BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic diagram of the dual nature of the stRNA and siRNA pathways.

FIG. 2A is a schematic representation of a wild-type, stRNA precursor (SEQ ID NO:1).

FIGs. 2B to 2E are schematic representations of synthetic, engineered RNA precursors (SEQ ID NOS:2, 3, 4, and 5).

FIG. 3 is an autoradiograph showing the results of an assay for determining whether an engineered RNA precursor can promote cleavage of the target mRNA *in vitro* in a standard RNAi reaction.

FIGs. 4A to 4C are schematic representations of synthetic luciferase siRNA (4A; SEQ ID NOS: 6 and 7), and 5' and 3' synthetic, engineered RNA precursors (4B; SEQ ID NO:8; and 4C; SEQ ID NO:9).

FIG. 4D is a schematic representation of a chimeric target mRNA for an *in vitro* luciferase/*let-7* RNAi reaction. The sites of siRNA-directed target cleavage are indicated by "scissors."

FIG. 4E is an autoradiograph showing the results of an assay for determining whether the 5' and 3' synthetic, engineered RNA precursors of FIGs. 4B and 4C can promote cleavage of the target mRNA *in vitro* in a standard RNAi reaction.

FIG. 5 is a schematic diagram of transgene encoding an engineered RNA precursor (SEQ ID NO:2) and the transcription and processing of the precursor to form a double-stranded siRNA (SEQ ID NO:7 and SEQ ID NO:12).



### DETAILED DESCRIPTION

Small temporal RNAs (stRNAs), also known as microRNAs (miRNAs), such as *lin-4* and *let-7* in *Caenorhabditis elegans* and *let-7* in *Drosophila melanogaster* and humans encode no protein, but instead appear to block the productive translation of mRNA by binding sequences in the 3' untranslated region (3' UTR) of their target mRNAs. As described in Hutvagner et al., Science, 293:834 (July 12, 2001), *let-7* RNA in *Drosophila* has been shown to be cleaved from a larger precursor transcript, which is similar to the generation of small RNAs from a longer, structured precursor double-stranded RNA in the RNA interference (RNAi) pathway.

Like siRNA, stRNAs are also 21-25 nucleotides long, but unlike siRNAs, are single-stranded and do not mediate gene silencing via target mRNA cleavage. As shown in FIG. 1, the RNAi and stRNA pathways intersect; both require the RNA processing enzyme Dicer to produce the active small RNA components that repress gene expression. Dicer and perhaps other proteins act on pre-stRNAs to yield mature, single-stranded stRNAs that repress mRNA translation. In RNAi, Dicer cleaves long, double-stranded RNA to yield siRNA duplexes that mediate targeted mRNA destruction.

Whereas long, double-stranded RNAs are cleaved symmetrically by Dicer to generate duplex siRNAs, current evidence suggests that stRNAs are cleaved asymmetrically to generate only a single-stranded stRNA. stRNA precursors are stem-loop RNAs that do not mediate target cleavage or provoke the sequence non-specific responses induced by long, double-stranded RNA. On the other hand, the invention provides new, engineered RNA precursors that when processed within a cell generate siRNAs that mediated target cleavage. These siRNAs can be double- or single-stranded, as long as they mediate cleavage of the target mRNA. Such engineered RNA precursors can be expressed in transgenic mammals in a cell-type-specific or developmental-stage-specific manner to induce RNAi in a specific cell or cells at a defined time.

A *Drosophila* embryo lysate that mediates RNAi in vitro (Tuschl et al., (1999) cited *supra*), which process double-stranded RNA into siRNA (Zamore et al., (2000) cited *supra*), and pre-*let-7*-stRNA into mature *let-7* stRNA (Hutvagner et al., (2001, cited *supra*), can be used to assay the ability of an engineered RNA precursor to

mediate RNAi *in vitro*. This assay allows testing of the new engineered RNA precursors. The new engineered precursors differ from naturally occurring, wild-type stRNA precursors by various modifications and by the fact that one portion of their duplex stem comprises a nucleic acid sequence that is complementary, preferably fully complementary, to a portion of the mRNA of a target gene.

#### Engineered RNA Precursors That Generate siRNAs

Naturally-occurring stRNA precursors (pre-stRNA) have certain elements or components that are illustrated in FIG. 2A, which shows an stRNA precursor for *let-7* (pre-*let-7*). Each precursor is a single strand that forms a duplex stem including two portions that are generally complementary, and a loop, that connects the two portions of the stem. In typical pre-stRNAs, the stem includes one or more bulges, e.g., extra nucleotides that create a single nucleotide "loop" in one portion of the stem, and/or one or more unpaired nucleotides that create a gap in the hybridization of the two portions of the stem to each other.

Engineered RNA precursors of the invention are artificial constructs that are similar to naturally occurring pre-stRNAs, but differ from the wild-type precursor sequences in a number of ways. The key difference is that one portion of the duplex stem is a nucleic acid sequence that is complementary (or anti-sense) to the target mRNA. Thus, engineered RNA precursors include a duplex stem with two portions and a loop connecting the two stem portions. The two stem portions are about 18 or 19 to about 25, 30, 35, 37, 38, 39, or 40 or more nucleotides in length. When used in mammalian cells, the length of the stem portions should be less than about 30 nucleotides to avoid provoking non-specific responses like the interferon pathway. In non-mammalian cells, the stem can be longer than 30 nucleotides. In fact, the stem can include much larger sections complementary to the target mRNA (up to, and including the entire mRNA). The two portions of the duplex stem must be sufficiently complementary to hybridize to form the duplex stem. Thus, the two portions can be, but need not be, fully or perfectly complementary. In addition, the two stem portions can be the same length, or one portion can include an overhang of 1, 2, 3, or 4 nucleotides. The overhanging nucleotides can include, for example, uracils (Us), e.g., all Us.

Other differences from natural pre-stRNA sequences include, but are not limited to, deleting unpaired or bulged nucleotides, introducing additional base-paired nucleotides to one or both of the stem portions, modifying the loop sequence to increase or decrease the number of paired nucleotides, or replacing all or part of the loop sequence with a tetraloop or other loop sequences. Thus, the loop in the engineered RNA precursors can be 2, 3, 4, 5, 6, 7, 8, 9, or more, e.g., 15 or 20, or more nucleotides in length. Tetraloop sequences can include, but are not limited to, the sequences GNRA (SEQ ID NO: 13), where N is any nucleotide and R is a purine nucleotide, GGGG (SEQ ID NO:14), and UUUU (SEQ ID NO:15).

Four examples of such engineered RNA precursors are illustrated in FIGs. 2B to 2E. FIGs. 2B and 2C illustrate engineered precursors in which the stem portions have had all unpaired and bulging nucleotides removed or paired, but the loop is the same as the wild-type loop in the pre-stRNA. FIGs. 2D and 2E illustrate two engineered RNA precursors with a tetraloop. In FIG 2D, the tetraloop UUUU (SEQ ID NO:15) replaces a portion of the wild-type loop in FIG 2A. In FIG 2E, the tetraloop GGGG (SEQ ID NO:14) replaces the entire wild-type loop sequence.

FIGs. 4B and 4C illustrate additional engineered RNA precursors. Each engineered RNA precursor includes in its stem a sequence that is perfectly complementary to a portion of the sequence of the firefly luciferase mRNA. In FIG 4B (SEQ ID NO:8), this region is shown in bold type, and is located on the 3' side of the stem. In FIG 4C (SEQ ID NO:9), this complementary sequence is on the 5' side of the stem. Unlike the naturally-occurring pre-let-7 RNA, these engineered RNA precursors have fully complementary stems, and direct RNAi against the luciferase mRNA.

In addition, modification of the naturally occurring stRNA precursor to generate an engineered RNA precursor (pre-siRNA) includes altering the sequence of the RNA to include the sequences of the desired siRNA duplex. The desired siRNA duplex, and thus both of the two stem portions in the engineered RNA precursor, are selected by methods known in the art. These include, but are not limited to, selecting an 18, 19, 20, 21 nucleotide, or longer, sequence from the target gene mRNA sequence from a region 100 to 200 or 300 nucleotides on the 3' side of the start of translation. In general, the sequence can be selected from any portion of the mRNA

from the target gene, such as the 5' UTR (untranslated region), coding sequence, or 3' UTR. This sequence can optionally follow immediately after a region of the target gene containing two adjacent AA nucleotides. The last two nucleotides of the 21 or so nucleotide sequence can be selected to be UU (so that the anti-sense strand of the siRNA begins with UU). This 21 or so nucleotide sequence is used to create one portion of a duplex stem in the engineered RNA precursor. This sequence can replace a stem portion of a wild-type pre-stRNA sequence, e.g., enzymatically, or is included in a complete sequence that is synthesized. For example, one can synthesize DNA oligonucleotides that encode the entire stem-loop engineered RNA precursor, or that encode just the portion to be inserted into the duplex stem of the precursor, and using restriction enzymes to build the engineered RNA precursor construct, e.g., from a wild-type pre-stRNA.

Engineered RNA precursors include in the duplex stem the 21-22 or so nucleotide sequences of the siRNA desired to be produced *in vivo*. Thus, the stem portion of the engineered RNA precursor includes at least 18 or 19 nucleotide pairs corresponding to the sequence of an exonic portion of the gene whose expression is to be reduced or inhibited. The two 3' nucleotides flanking this region of the stem are chosen so as to maximize the production of the siRNA from the engineered RNA precursor, and to maximize the efficacy of the resulting siRNA in targeting the corresponding mRNA for destruction by RNAi *in vivo* and *in vitro*.

Another defining feature of these engineered RNA precursors is that as a consequence of their length, sequence, and/or structure, they do not induce sequence non-specific responses, such as induction of the interferon response or apoptosis, or that they induce a lower level of such sequence non-specific responses than long, double-stranded RNA (>150 bp) currently used to induce RNAi. For example, the interferon response is triggered by dsRNA longer than 30 base pairs.

#### Transgenes Encoding Engineered RNA Precursors

The new engineered RNA precursors can be synthesized by standard methods known in the art, e.g., by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). These synthetic,

engineered RNA precursors can be used directly as described below or cloned into expression vectors by methods known in the field.

The engineered RNA precursors should be delivered to cells *in vitro* or *in vivo* in which it is desired to target a specific mRNA for destruction. A number of methods have been developed for delivering DNA or RNA to cells. For example, for *in vivo* delivery, molecules can be injected directly into a tissue site or administered systemically. *In vitro* delivery includes methods known in the art such as electroporation and lipofection.

To achieve intracellular concentrations of the nucleic acid molecule sufficient to suppress expression of endogenous mRNAs, one can use, for example, a recombinant DNA construct in which the oligonucleotide is placed under the control of a strong *Pol III* (e.g., U6 or *Pol III* H1-RNA promoter) or *Pol II* promoter. The use of such a construct to transfect target cells *in vitro* or *in vivo* will result in the transcription of sufficient amounts of the engineered RNA precursor to lead to the production of an siRNA that can target a corresponding mRNA sequence for cleavage by RNAi to decrease the expression of the gene encoding that mRNA. For example, a vector can be introduced *in vivo* such that it is taken up by a cell and directs the transcription of an engineered RNA precursor. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired siRNA precursor.

Such vectors can be constructed by recombinant DNA technology methods known in the art. Vectors can be plasmid, viral, or other vectors known in the art such as those described herein, used for replication and expression in mammalian cells or other targeted cell types. The nucleic acid sequences encoding the engineered RNA precursors can be prepared using known techniques. For example, two synthetic DNA oligonucleotides can be synthesized to create a novel gene encoding the entire engineered RNA precursor. The DNA oligonucleotides, which will pair, leaving appropriate 'sticky ends' for cloning, can be inserted into a restriction site in a plasmid that contains a promoter sequence (e.g., a *Pol II* or a *Pol III* promoter) and appropriate terminator sequences 3' to the engineered RNA precursor sequences (e.g., a cleavage and polyadenylation signal sequence from SV40 or a *Pol III* terminator sequence).

The invention also encompasses genetically engineered host cells that contain any of the foregoing expression vectors and thereby express the nucleic acid molecules of the invention in the host cell. The host cells can be cultured using known techniques and

methods (see, e.g., Culture of Animal Cells (R.I. Freshney, Alan R. Liss, Inc. 1987); Molecular Cloning, Sambrook et al. (Cold Spring Harbor Laboratory Press, 1989)).

Successful introduction of the vectors of the invention into host cells can be monitored using various known methods. For example, transient transfection can be signaled with a reporter, such as a fluorescent marker, such as Green Fluorescent Protein (GFP). Stable transfection can be indicated using markers that provide the transfected cell with resistance to specific environmental factors (e.g., antibiotics and drugs), such as hygromycin B resistance, e.g., in insect cells and in mammalian cells.

#### 10 Regulatory Sequences

The expression of the engineered RNA precursors is driven by regulatory sequences, and the vectors of the invention can include any regulatory sequences known in the art to act in mammalian cells, e.g., murine cells; in insect cells; in plant cells; or other cells. The term regulatory sequence includes promoters, enhancers, and other expression control elements. It will be appreciated that the appropriate regulatory sequence depends on such factors as the future use of the cell or transgenic animal into which a sequence encoding an engineered RNA precursor is being introduced, and the level of expression of the desired RNA precursor. A person skilled in the art would be able to choose the appropriate regulatory sequence. For example, the transgenic animals described herein can be used to determine the role of a test polypeptide or the engineered RNA precursors in a particular cell type, e.g., a hematopoietic cell. In this case, a regulatory sequence that drives expression of the transgene ubiquitously, or a hematopoietic-specific regulatory sequence that expresses the transgene only in hematopoietic cells, can be used. Expression of the engineered RNA precursors in a hematopoietic cell means that the cell is now susceptible to specific, targeted RNAi of a particular gene. Examples of various regulatory sequences are described below.

The regulatory sequences can be inducible or constitutive. Suitable constitutive regulatory sequences include the regulatory sequence of a housekeeping gene such as the  $\alpha$ -actin regulatory sequence, or may be of viral origin such as regulatory sequences derived from mouse mammary tumor virus (MMTV) or cytomegalovirus (CMV).

Alternatively, the regulatory sequence can direct transgene expression in specific organs or cell types (see, e.g., Lasko et al., 1992, Proc. Natl. Acad. Sci. USA 89:6232). Several tissue-specific regulatory sequences are known in the art including the albumin regulatory sequence for liver (Pinkert et al., 1987, Genes Dev. 1:268-276); the endothelin regulatory sequence for endothelial cells (Lee, 1990, J. Biol. Chem. 265:10446-50); the keratin regulatory sequence for epidermis; the myosin light chain-2 regulatory sequence for heart (Lee et al., 1992, J. Biol. Chem. 267:15875-85), and the insulin regulatory sequence for pancreas (Bucchini et al., 1986, Proc. Natl. Acad. Sci. USA 83:2511-2515), or the vav regulatory sequence for hematopoietic cells (Oligvy et al., 1999, Proc. Natl. Acad. Sci. USA 96:14943-14948). Another suitable regulatory sequence, which directs constitutive expression of transgenes in cells of hematopoietic origin, is the murine MHC class I regulatory sequence (Morello et al., 1986, EMBO J. 5:1877-1882). Since MHC expression is induced by cytokines, expression of a test gene operably linked to this regulatory sequence can be upregulated in the presence of cytokines.

In addition, expression of the transgene can be precisely regulated, for example, by using an inducible regulatory sequence and expression systems such as a regulatory sequence that is sensitive to certain physiological regulators, e.g., circulating glucose levels, or hormones (Docherty et al., 1994, FASEB J. 8:20-24). Such inducible expression systems, suitable for the control of transgene expression in cells or in mammals such as mice, include regulation by ecdysone, by estrogen, progesterone, tetracycline, chemical inducers of dimerization, and isopropyl-beta-D-1-thiogalactopyranoside (IPTG)(collectively referred to as "the regulatory molecule"). Each of these expression systems is well described in the literature and permits expression of the transgene throughout the animal in a manner controlled by the presence or absence of the regulatory molecule. For a review of inducible expression systems, see, e.g., Mills, 2001, Genes Devel. 15:1461-1467, and references cited therein.

The regulatory elements referred to above include, but are not limited to, the cytomegalovirus hCMV immediate early gene, the early or late promoters of SV40 adenovirus (Bernoist et al., *Nature*, 290:304, 1981), the tet system, the lac system, the trp system, the TAC system, the TRC system, the major operator and promoter regions of

phage A, the control regions of fd coat protein, the promoter for 3-phosphoglycerate kinase, the promoters of acid phosphatase, and the promoters of the yeast  $\alpha$ -mating factors. Additional promoters include the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., *Cell* 22:787-797, 1988); the herpes  
5 thymidine kinase promoter (Wagner et al., *Proc. Natl. Acad. Sci. USA* 78:1441, 1981); or the regulatory sequences of the metallothionein gene (Brinster et al., *Nature* 296:39, 1988).

#### Assay for Testing Engineered RNA Precursors

10 *Drosophila* embryo lysates can be used to determine if an engineered RNA precursor was, in fact, the direct precursor of a mature stRNA or siRNA. This lysate assay is described in Tuschl et al., 1999, *supra*, Zamore et al., 2000, *supra*, and Hutvagner et al. 2001, *supra*. These lysates recapitulate RNAi *in vitro*, thus  
15 permitting investigation into whether the proposed precursor RNA was cleaved into a mature stRNA or siRNA by an RNAi-like mechanism. Briefly, the precursor RNA is incubated with *Drosophila* embryo lysate for various times, then assayed for the production of the mature siRNA or stRNA by primer extension or Northern hybridization. As in the *in vivo* setting, mature RNA accumulates in the cell-free reaction. Thus, an RNA corresponding to the proposed precursor can be shown to be  
20 converted into a mature stRNA or siRNA duplex in the *Drosophila* embryo lysate.

Furthermore, an engineered RNA precursor can be functionally tested in the *Drosophila* embryo lysates. In this case, the engineered RNA precursor is incubated in the lysate in the presence of a 5' radiolabeled target mRNA in a standard *in vitro* RNAi reaction for various lengths of time. The target mRNA can be 5' radiolabeled  
25 using guanylyl transferase (as described in Tuschl et al., 1999, *supra* and references therein) or other suitable methods. The products of the *in vitro* reaction are then isolated and analyzed on a denaturing acrylamide or agarose gel to determine if the target mRNA has been cleaved in response to the presence of the engineered RNA precursor in the reaction. The extent and position of such cleavage of the mRNA  
30 target will indicate if the engineering of the precursor created a pre-siRNA capable of mediating sequence-specific RNAi.



### Transgenic Animals

Engineered RNA precursors of the invention can be expressed in transgenic animals. These animals represent a model system for the study of disorders that are caused by, or exacerbated by, overexpression or underexpression (as compared to wild-type or normal) of nucleic acids (and their encoded polypeptides) targeted for destruction by the engineered RNA precursor products (siRNAs), and for the development of therapeutic agents that modulate the expression or activity of nucleic acids or polypeptides targeted for destruction.

Transgenic animals can be farm animals (pigs, goats, sheep, cows, horses, rabbits, and the like), rodents (such as rats, guinea pigs, and mice), non-human primates (for example, baboons, monkeys, and chimpanzees), and domestic animals (for example, dogs and cats). Invertebrates such as *Caenorhabditis elegans* or *Drosophila* can be used as well as non-mammalian vertebrates such as fish (e.g., zebrafish) or birds (e.g., chickens). Engineered RNA precursors with stems of 18 to 30 nucleotides in length are preferred for use in mammals, such as mice.

A transgenic founder animal can be identified based upon the presence of a transgene that encodes the new RNA precursors in its genome, and/or expression of the transgene in tissues or cells of the animals, for example, using PCR or Northern analysis. Expression is confirmed by a decrease in the expression (RNA or protein) of the target sequence.

A transgenic founder animal can be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding the RNA precursors can further be bred to other transgenic animals carrying other transgenes. In addition, cells obtained from the transgenic founder animal or its offspring can be cultured to establish primary, secondary, or immortal cell lines containing the transgene.

### Procedures for Making Transgenic, Non-Human Animals

A number of methods have been used to obtain transgenic, non-human animals, which are animals that have gained an additional gene by the introduction of a transgene into their cells (e.g., both the somatic and germ cells), or into an ancestor's germ line. In some cases, transgenic animals can be generated by commercial facilities (e.g., The Transgenic *Drosophila* Facility at Michigan State University, The

Transgenic Zebrafish Core Facility at the Medical College of Georgia (Augusta, Georgia), and Xenogen Biosciences (St. Louis, MO). In general, the construct containing the transgene is supplied to the facility for generating a transgenic animal.

Methods for generating transgenic animals include introducing the transgene  
5 into the germ line of the animal. One method is by microinjection of a gene construct into the pronucleus of an early stage embryo (e.g., before the four-cell stage; Wagner et al., 1981, Proc. Natl. Acad. Sci. USA 78:5016; Brinster et al., 1985, Proc. Natl. Acad. Sci. USA 82:4438). Alternatively, the transgene can be introduced into the pronucleus by retroviral infection. A detailed procedure for producing such  
10 transgenic mice has been described (see e.g., Hogan et al., Manipulating the Mouse Embryo, Cold Spring Harbour Laboratory, Cold Spring Harbour, NY (1986); U.S. Patent No. 5,175,383 (1992)). This procedure has also been adapted for other animal species (e.g., Hammer et al., 1985, Nature 315:680; Murray et al., 1989, Reprod. Fert. Devl. 1:147; Pursel et al., 1987, Vet. Immunol. Histopath. 17:303; Rexroad et al.,  
15 1990, J. Reprod. Fert. 41 (suppl):119; Rexroad et al., 1989, Molec. Reprod. Devl. 1:164; Simons et al., 1988, BioTechnology 6:179; Vize et al., 1988, J. Cell. Sci. 90:295; and Wagner, 1989, J. Cell. Biochem. 13B (suppl):164).

In brief, the procedure involves introducing the transgene into an animal by microinjecting the construct into the pronuclei of the fertilized mammalian egg(s) to  
20 cause one or more copies of the transgene to be retained in the cells of the developing mammal(s). Following introduction of the transgene construct into the fertilized egg, the egg may be incubated *in vitro* for varying amounts of time, or reimplanted in a surrogate host, or both. One common method is to incubate the embryos *in vitro* for about 1-7 days, depending on the species, and then reimplant them into the surrogate  
25 host. The presence of the transgene in the progeny of the transgenically manipulated embryos can be tested by Southern blot analysis of a segment of tissue.

Another method for producing germ-line transgenic animals is through the use of embryonic stem (ES) cells. The gene construct can be introduced into embryonic stem cells by homologous recombination (Thomas et al., 1987, Cell 51:503;  
30 Capecchi, Science 1989, 244:1288; Joyner et al., 1989, Nature 338:153) in a transcriptionally active region of the genome. A suitable construct can also be introduced into embryonic stem cells by DNA-mediated transfection, such as by

electroporation (Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, 1987). Detailed procedures for culturing embryonic stem cells (e.g., ES-D3, ATCC# CCL-1934, ES-E14TG2a, ATCC# CCL-1821, American Type Culture Collection, Rockville, MD) and methods of making transgenic animals from  
5 embryonic stem cells can be found in Teratocarcinomas and Embryonic Stem Cells, A Practical Approach, ed. E.J. Robertson (IRL Press, 1987). In brief, the ES cells are obtained from pre-implantation embryos cultured *in vitro* (Evans et al., 1981, Nature 292:154-156). Transgenes can be efficiently introduced into ES cells by DNA  
10 transfection or by retrovirus-mediated transduction. The resulting transformed ES cells can thereafter be combined with blastocysts from a non-human animal. The ES cells colonize the embryo and contribute to the germ line of the resulting chimeric animal.

In the above methods, the transgene can be introduced as a linear construct, a circular plasmid, or a viral vector, which can be incorporated and inherited as a  
15 transgene integrated into the host genome. The transgene can also be constructed to permit it to be inherited as an extrachromosomal plasmid (Gassmann et al., 1995, Proc. Natl. Acad. Sci. USA 92:1292). A plasmid is a DNA molecule that can replicate autonomously in a host.

The transgenic, non-human animals can also be obtained by infecting or  
20 transfecting cells either *in vivo* (e.g., direct injection), *ex vivo* (e.g., infecting the cells outside the host and later reimplanting), or *in vitro* (e.g., infecting the cells outside host), for example, with a recombinant viral vector carrying a gene encoding the engineered RNA precursors. Examples of suitable viral vectors include recombinant retroviral vectors (Valerio et al., 1989, Gene 84:419; Scharfman et al., 1991, Proc.  
25 Natl. Acad. Sci. USA 88:462; Miller and Buttimore, 1986, Mol. Cell. Biol. 6:2895), recombinant adenoviral vectors (Freidman et al., 1986, Mol. Cell. Biol. 6:3791; Levrero et al., 1991, Gene 101:195), and recombinant Herpes simplex viral vectors (Fink et al., 1992, Human Gene Therapy 3:11). Such methods are also useful for introducing constructs into cells for uses other than generation of transgenic animals.

30 Other approaches include insertion of transgenes encoding the new engineered RNA precursors into viral vectors including recombinant adenovirus, adeno-associated virus, and herpes simplex virus-1, or recombinant bacterial or eukaryotic

plasmids. Viral vectors transfect cells directly. Other approaches include delivering the transgenes, in the form of plasmid DNA, with the help of, for example, cationic liposomes (lipofectin) or derivatized (e.g. antibody conjugated) polylysine conjugates, gramacidin S, artificial viral envelopes, or other such intracellular carriers, as well as  
5 direct injection of the transgene construct or CaPO<sub>4</sub> precipitation carried out *in vivo*. Such methods can also be used *in vitro* to introduce constructs into cells for uses other than generation of transgenic animals.

Retrovirus vectors and adeno-associated virus vectors can be used as a recombinant gene delivery system for the transfer of exogenous genes *in vivo* or *in*  
10 *vitro*. These vectors provide efficient delivery of genes into cells, and the transferred nucleic acids are stably integrated into the chromosomal DNA of the host. The development of specialized cell lines (termed "packaging cells") which produce only replication-defective retroviruses has increased the utility of retroviruses for gene therapy, and defective retroviruses are characterized for use in gene transfer for gene  
15 therapy purposes (for a review see Miller, 1990, Blood 76:271). A replication-defective retrovirus can be packaged into virions which can be used to infect a target cell through the use of a helper virus by standard techniques. Protocols for producing recombinant retroviruses and for infecting cells *in vitro* or *in vivo* with such viruses can be found in Current Protocols in Molecular Biology, Ausubel, F.M. et al. (eds.)  
20 Greene Publishing Associates, (1989), Sections 9.10-9.14 and other standard laboratory manuals.

Examples of suitable retroviruses include pLJ, pZIP, pWE and pEM which are known to those skilled in the art. Examples of suitable packaging virus lines for preparing both ecotropic and amphotropic retroviral systems include Psi-Crip, Psi-  
25 Cre, Psi-2 and Psi-Am. Retroviruses have been used to introduce a variety of genes into many different cell types, including epithelial cells, *in vitro* and/or *in vivo* (see for example Eglitis, et al., 1985, Science 230:1395-1398; Danos and Mulligan, 1988, Proc. Natl. Acad. Sci. USA 85:6460-6464; Wilson et al., 1988, Proc. Natl. Acad. Sci. USA 85:3014-3018; Armentano et al., 1990, Proc. Natl. Acad. Sci. USA 87:6141-  
30 6145; Huber et al., 1991, Proc. Natl. Acad. Sci. USA 88:8039-8043; Ferry et al., 1991, Proc. Natl. Acad. Sci. USA 88:8377-8381; Chowdhury et al., 1991, Science 254:1802-1805; van Beusechem et al., 1992, Proc. Natl. Acad. Sci. USA 89:7640-

7644; Kay et al., 1992, Human Gene Therapy 3:641-647; Dai et al., 1992, Proc. Natl. Acad. Sci. USA 89:10892-10895; Hwu et al., 1993, J. Immunol. 150:4104-4115; U.S. Patent No. 4,868,116; U.S. Patent No. 4,980,286; PCT Application WO 89/07136; PCT Application WO 89/02468; PCT Application WO 89/05345; and PCT Application WO 92/07573).

In another example, recombinant retroviral vectors capable of transducing and expressing genes inserted into the genome of a cell can be produced by transfecting the recombinant retroviral genome into suitable packaging cell lines such as PA317 and Psi-CRIP (Cornette et al., 1991, Human Gene Therapy 2:5-10; Cone et al., 1984, Proc. Natl. Acad. Sci. USA 81:6349). Recombinant adenoviral vectors can be used to infect a wide variety of cells and tissues in susceptible hosts (e.g., rat, hamster, dog, and chimpanzee) (Hsu et al., 1992, J. Infectious Disease, 166:769), and also have the advantage of not requiring mitotically active cells for infection.

Another viral gene delivery system useful in the present invention also utilizes adenovirus-derived vectors. The genome of an adenovirus can be manipulated such that it encodes and expresses a gene product of interest but is inactivated in terms of its ability to replicate in a normal lytic viral life cycle. See, for example, Berkner et al. (1988, BioTechniques 6:616), Rosenfeld et al. (1991, Science 252:431-434), and Rosenfeld et al. (1992, Cell 68:143-155). Suitable adenoviral vectors derived from the adenovirus strain Ad type 5 dl324 or other strains of adenovirus (e.g., Ad2, Ad3, Ad7 etc.) are known to those skilled in the art. Recombinant adenoviruses can be advantageous in certain circumstances in that they are not capable of infecting nondividing cells and can be used to infect a wide variety of cell types, including epithelial cells (Rosenfeld et al., 1992, cited *supra*). Furthermore, the virus particle is relatively stable and amenable to purification and concentration, and as above, can be modified to affect the spectrum of infectivity. Additionally, introduced adenoviral DNA (and foreign DNA contained therein) is not integrated into the genome of a host cell but remains episomal, thereby avoiding potential problems that can occur as a result of insertional mutagenesis *in situ* where introduced DNA becomes integrated into the host genome (e.g., retroviral DNA). Moreover, the carrying capacity of the adenoviral genome for foreign DNA is large (up to 8 kilobases) relative to other gene

delivery vectors (Berkner et al. cited *supra*; Haj-Ahmand and Graham, 1986, *J. Virol.* 57:267).

Yet another viral vector system useful for delivery of the subject transgenes is the adeno-associated virus (AAV). Adeno-associated virus is a naturally occurring defective virus that requires another virus, such as an adenovirus or a herpes virus, as a helper virus for efficient replication and a productive life cycle. For a review, see Muzyczka et al. (1992, *Curr. Topics in Micro. and Immunol.* 158:97-129). It is also one of the few viruses that may integrate its DNA into non-dividing cells, and exhibits a high frequency of stable integration (see for example Flotte et al. (1992, *Am. J. Respir. Cell. Mol. Biol.* 7:349-356; Samulski et al., 1989, *J. Virol.* 63:3822-3828; and McLaughlin et al. (1989, *J. Virol.* 62:1963-1973). Vectors containing as little as 300 base pairs of AAV can be packaged and can integrate. Space for exogenous DNA is limited to about 4.5 kb. An AAV vector such as that described in Tratschin et al. (1985) *Mol. Cell. Biol.* 5:3251-3260 can be used to introduce DNA into cells. A variety of nucleic acids have been introduced into different cell types using AAV vectors (see for example Hermonat et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6466-6470; Tratschin et al. (1985) *Mol. Cell. Biol.* 4:2072-2081; Wondisford et al. (1988) *Mol. Endocrinol.* 2:32-39; Tratschin et al. (1984) *J. Virol.* 51:611-619; and Flotte et al. (1993) *J. Biol. Chem.* 268:3781-3790).

In addition to viral transfer methods, such as those illustrated above, non-viral methods can also be employed to cause expression of an engineered RNA precursor of the invention in the tissue of an animal. Most non-viral methods of gene transfer rely on normal mechanisms used by mammalian cells for the uptake and intracellular transport of macromolecules. In preferred embodiments, non-viral gene delivery systems of the present invention rely on endocytic pathways for the uptake of the subject gene of the invention by the targeted cell. Exemplary gene delivery systems of this type include liposomal derived systems, poly-lysine conjugates, and artificial viral envelopes. Other embodiments include plasmid injection systems such as are described in Meuli et al., (2001) *J. Invest. Dermatol.*, 116(1):131-135; Cohen et al., (2000) *Gene Ther.*, 7(22):1896-905; and Tam et al., (2000) *Gene Ther.*, 7(21):1867-74.

In a representative embodiment, a gene encoding an engineered RNA precursor of the invention can be entrapped in liposomes bearing positive charges on their surface (e.g., lipofectins) and (optionally) which are tagged with antibodies against cell surface antigens of the target tissue (Mizuno et al., (1992) *No Shinkei Geka*, 20:547-551; PCT publication WO91/06309; Japanese patent application 1047381; and European patent publication EP-A-43075).

Animals harboring the transgene can be identified by detecting the presence of the transgene in genomic DNA (e.g., using Southern analysis). In addition, expression of the engineered RNA precursor can be detected directly (e.g., by Northern analysis). Expression of the transgene can also be confirmed by detecting a decrease in the amount of protein corresponding to the targeted sequence. When the transgene is under the control of an inducible or developmentally regulated promoter, expression of the target protein is decreased when the transgene is induced or at the developmental stage when the transgene is expressed, respectively.

#### Clones of Transgenic Animals

Clones of the non-human transgenic animals described herein can be produced according to the methods described in Wilmut et al. ((1997) *Nature*, 385:810-813) and PCT publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, e.g., a somatic cell from the transgenic animal, can be isolated and induced to exit the growth cycle and enter the G<sub>0</sub> phase to become quiescent. The quiescent cell can then be fused, e.g., through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops into a morula or blastocyte and is then transferred to a pseudopregnant female foster animal. Offspring borne of this female foster animal will be clones of the animal from which the cell, e.g., the somatic cell, was isolated.

Once the transgenic animal is produced, cells of the transgenic animal and cells from a control animal are screened to determine the presence of an RNA precursor nucleic acid sequence, e.g., using polymerase chain reaction (PCR). Alternatively, the cells can be screened to determine if the RNA precursor is expressed (e.g., by standard procedures such as Northern blot analysis or reverse

transcriptase-polymerase chain reaction (RT-PCR); Sambrook et al., *Molecular Cloning - A Laboratory Manual*, (Cold Spring Harbor Laboratory, 1989)).

The transgenic animals of the present invention can be homozygous or heterozygous, and one of the benefits of the invention is that the target mRNA is effectively degraded even in heterozygotes. The present invention provides for transgenic animals that carry a transgene of the invention in all their cells, as well as animals that carry a transgene in some, but not all of their cells. That is, the invention provides for mosaic animals. The transgene can be integrated as a single transgene or in concatamers, e.g., head-to-head tandems or head-to-tail tandems.

For a review of techniques that can be used to generate and assess transgenic animals, skilled artisans can consult Gordon (*Intl. Rev. Cytol.* 115:171-229, 1989), and may obtain additional guidance from, for example: Hogan et al. "Manipulating the Mouse Embryo" (Cold Spring Harbor Press, Cold Spring Harbor, NY, 1986; Krimpenfort et al., *Bio/Technology* 9:86, 1991; Palmiter et al., *Cell* 41:343, 1985; Kraemer et al., "Genetic Manipulation of the Early Mammalian Embryo," Cold Spring Harbor Press, Cold Spring Harbor, NY, 1985; Hammer et al., *Nature* 315:680, 1985; Purcel et al., *Science*, 244:1281, 1986; Wagner et al., U.S. Patent No. 5,175,385; and Krimpenfort et al., U.S. Patent No. 5,175,384.

### Transgenic Plants

Among the eukaryotic organisms featured in the invention are plants containing an exogenous nucleic acid that encodes an engineered RNA precursor of the invention.

Accordingly, a method according to the invention comprises making a plant having a nucleic acid molecule or construct, e.g., a transgene, described herein. Techniques for introducing exogenous nucleic acids into monocotyledonous and dicotyledonous plants are known in the art, and include, without limitation, *Agrobacterium*-mediated transformation, viral vector-mediated transformation, electroporation and particle gun transformation, see, e.g., U.S. Patents Nos. 5,204,253 and 6,013,863. If a cell or tissue culture is used as the recipient tissue for



transformation, plants can be regenerated from transformed cultures by techniques known to those skilled in the art. Transgenic plants can be entered into a breeding program, *e.g.*, to introduce a nucleic acid encoding a polypeptide into other lines, to transfer the nucleic acid to other species or for further selection of other desirable traits. Alternatively, transgenic plants can be propagated vegetatively for those species amenable to such techniques. Progeny includes descendants of a particular plant or plant line. Progeny of a plant include seeds formed on F<sub>1</sub>, F<sub>2</sub>, F<sub>3</sub>, and subsequent generation plants, or seeds formed on BC<sub>1</sub>, BC<sub>2</sub>, BC<sub>3</sub>, and subsequent generation plants. Seeds produced by a transgenic plant can be grown and then selfed (or outcrossed and selfed) to obtain seeds homozygous for the nucleic acid encoding a novel polypeptide.

A suitable group of plants with which to practice the invention include dicots, such as safflower, alfalfa, soybean, rapeseed (high erucic acid and canola), or sunflower. Also suitable are monocots such as corn, wheat, rye, barley, oat, rice, millet, amaranth or sorghum. Also suitable are vegetable crops or root crops such as potato, broccoli, peas, sweet corn, popcorn, tomato, beans (including kidney beans, lima beans, dry beans, green beans) and the like. Also suitable are fruit crops such as peach, pear, apple, cherry, orange, lemon, grapefruit, plum, mango and palm. Thus, the invention has use over a broad range of plants, including species from the genera *Anacardium*, *Arachis*, *Asparagus*, *Atropa*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Carthamus*, *Cocos*, *Coffea*, *Cucumis*, *Cucurbita*, *Daucus*, *Elaeis*, *Fragaria*, *Glycine*, *Gossypium*, *Helianthus*, *Heterocallis*, *Hordeum*, *Hyoscyamus*, *Lactuca*, *Linum*, *Lolium*, *Lupinus*, *Lycopersicon*, *Malus*, *Manihot*, *Majorana*, *Medicago*, *Nicotiana*, *Olea*, *Oryza*, *Panicum*, *Pannasetum*, *Persea*, *Phaseolus*, *Pistachia*, *Pisum*, *Pyrus*, *Prunus*, *Raphanus*, *Ricinus*, *Secale*, *Senecio*, *Sinapis*, *Solanum*, *Sorghum*, *Theobromus*, *Trigonella*, *Triticum*, *Vicia*, *Vitis*, *Vigna* and *Zea*.

The nucleic acid molecules of the invention can be expressed in plants in a cell- or tissue-specific manner according to the regulatory elements chosen to include in a particular nucleic acid construct present in the plant. Suitable cells, tissues, and organs in which to express a chimeric polypeptide of the invention include, without limitation, egg cell, central cell, synergid cell, zygote, ovule primordia, nucellus, integuments, endothelium, female gametophyte cells, embryo, axis, cotyledons,

suspensor, endosperm, seed coat, ground meristem, vascular bundle, cambium, phloem, cortex, shoot or root apical meristems, lateral shoot or root meristems, floral meristem, leaf primordia, leaf mesophyll cells, and leaf epidermal cells, e.g., epidermal cells involved in forming the cuticular layer. Also suitable are cells and  
5 tissues grown in liquid media or on semi-solid media.

### Transgenic Fungi

Other eukaryotic organisms featured in the invention are fungi containing an exogenous nucleic acid molecule that encodes an engineered RNA precursor of the  
10 invention.

Accordingly, a method according to the invention comprises introducing a nucleic acid molecule or construct as described herein into a fungus. Techniques for introducing exogenous nucleic acids into many fungi are known in the art, see, e.g., U.S. Patents Nos. 5,252,726 and 5,070,020. Transformed fungi can be cultured by  
15 techniques known to those skilled in the art. Such fungi can be used to introduce a nucleic acid encoding a polypeptide into other fungal strains, to transfer the nucleic acid to other species or for further selection of other desirable traits.

A suitable group of fungi with which to practice the invention include fission yeast and budding yeast, such as *Saccharomyces cerevisiae*, *S. pombe*, *S.*  
20 *carlsbergeris* and *Candida albicans*. Filamentous fungi such as *Aspergillus spp.* and *Penicillium spp.* are also useful.

### Pharmaceutical Compositions

The molecules of the invention can be incorporated into pharmaceutical compositions. Such compositions typically include a nucleic acid molecule, e.g., a  
25 nucleic acid molecule (e.g., a transgene) that encodes an engineered RNA precursor, or the precursor RNA itself, and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" includes solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical  
30 administration. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, inhalation, transdermal (topical), transmucosal, and rectal administration. Administration can also be oral.

5 Solutions or suspensions used for parenteral administration such as intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; 10 chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes, or multiple dose vials made of glass or plastic.

15 Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered 20 saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It should be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, 25 propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, 30 chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the

injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound in the required amount in an appropriate solvent with one or a combination  
5 of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are  
10 vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules, e.g.,  
15 gelatin capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin;  
20 an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel™, or corn starch; a lubricant such as magnesium stearate or Sterotes™; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

25 For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier  
30 to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished

through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

5 The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems.  
10 Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate; polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to  
15 infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is advantageous to formulate oral or parenteral compositions in dosage unit  
20 form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier.

25 Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD<sub>50</sub> (the dose lethal to 50% of the population) and the ED<sub>50</sub> (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the  
30 ratio LD<sub>50</sub>/ED<sub>50</sub>. Compounds which exhibit high therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in

order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of compositions of the invention lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range of the compound or, when appropriate, of the polypeptide product of a target sequence (e.g., achieving a decreased concentration of the polypeptide) that includes the IC50 (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

A therapeutically effective amount of a composition containing a sequence that encodes an engineered RNA precursor, or the precursor itself (i.e., an effective dosage), is an amount that inhibits expression of the polypeptide encoded by the target gene by at least 30 percent. Higher percentages of inhibition, e.g., 45, 50, 75, 85, 90 percent or higher may be preferred in certain embodiments. Exemplary doses include milligram or microgram amounts of the molecule per kilogram of subject or sample weight (e.g., about 1 microgram per kilogram to about 500 milligrams per kilogram, about 100 micrograms per kilogram to about 5 milligrams per kilogram, or about 1 microgram per kilogram to about 50 micrograms per kilogram. The compositions can be administered one time per week for between about 1 to 10 weeks, e.g., between 2 to 8 weeks, or between about 3 to 7 weeks, or for about 4, 5, or 6 weeks. The skilled artisan will appreciate that certain factors may influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and other diseases present. Moreover, treatment of a subject with a therapeutically effective amount of a composition can include a single treatment or a

series of treatments. In some cases transient expression of the engineered RNA precursor may be desired. When an inducible promoter is included in the construct encoding an engineered RNA precursor, expression is assayed upon delivery to the subject of an appropriate dose of the substance used to induce expression.

5           It is furthermore understood that appropriate doses of a composition depend upon the potency of the molecule (the sequence encoding the engineered precursor) with respect to the expression or activity to be modulated. When one or more of these molecules is to be administered to an animal (e.g., a human) to modulate expression or activity of a polypeptide or nucleic acid of the invention, a physician, veterinarian,  
10           or researcher may, for example, prescribe a relatively low dose at first, subsequently increasing the dose until an appropriate response is obtained. In addition, it is understood that the specific dose level for any particular subject will depend upon a variety of factors including the activity of the specific compound employed, the age, body weight, general health, gender, and diet of the subject, the time of  
15           administration, the route of administration, the rate of excretion, any drug combination, and the degree of expression or activity to be modulated.

          The nucleic acid molecules of the invention can be generally inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent  
20           5,328,470) or by stereotactic injection (see e.g., Chen et al. (1994) Proc. Natl. Acad. Sci. USA 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g.,  
25           retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

          The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

#### Methods of Treatment

30           The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder

associated with aberrant or unwanted expression or activity of any gene that is transcribed. As used herein, the term "treatment" is defined as the application or administration of a therapeutic agent to a patient, or application or administration of a therapeutic agent to an isolated tissue or cell line from a patient, who has a disorder, e.g., a disease or condition, a symptom of disease, or a predisposition toward a disease, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve, or affect the disease, the symptoms of disease, or the predisposition toward disease. A therapeutic agent is an engineered RNA precursor of the invention, or a nucleic acid molecule (DNA) that encodes the precursor.

With regards to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained about a subject's genome, specifically knowledge about a gene sequence (i.e., mutated gene) whose expression is associated with disease. Thus, a molecule of the invention can be engineered, based on knowledge of the gene whose expression is targeted, to inhibit expression of that gene as described herein.

Thus, in one aspect, the invention provides a method for treating in a subject, a disorder, e.g., a disease or condition, associated with an aberrant or unwanted gene expression or activity, by administering to the subject an engineered nucleic acid sequence that encodes an engineered precursor RNA. Subjects at risk for a disorder which is caused or contributed to by aberrant or unwanted expression or activity of a gene can be identified by, for example, any or a combination of diagnostic or prognostic assays that are known in the art. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the aberrance, such that a disease or disorder is prevented or, alternatively, delayed in its progression.

The molecules of the invention can act as novel therapeutic agents for controlling one or more of cellular proliferative and/or differentiative disorders, disorders associated with bone metabolism, immune disorders, hematopoietic disorders, cardiovascular disorders, liver disorders, viral diseases, pain or metabolic disorders.

Examples of cellular proliferative and/or differentiative disorders include cancer, e.g., carcinoma, sarcoma, metastatic disorders or hematopoietic neoplastic disorders, e.g., leukemias. A metastatic tumor can arise from a multitude of primary



tumor types, including but not limited to those of prostate, colon, lung, breast and liver origin.

As used herein, the terms "cancer," "hyperproliferative," and "neoplastic" refer to cells having the capacity for autonomous growth, i.e., an abnormal state or condition characterized by rapidly proliferating cell growth. Hyperproliferative and neoplastic disease states may be categorized as pathologic, i.e., characterizing or constituting a disease state, or may be categorized as non-pathologic, i.e., a deviation from normal but not associated with a disease state. The term is meant to include all types of cancerous growths or oncogenic processes, metastatic tissues or malignantly transformed cells, tissues, or organs, irrespective of histopathologic type or stage of invasiveness. "Pathologic hyperproliferative" cells occur in disease states characterized by malignant tumor growth. Examples of non-pathologic hyperproliferative cells include proliferation of cells associated with wound repair.

The terms "cancer" or "neoplasms" include malignancies of the various organ systems, such as affecting lung, breast, thyroid, lymphoid, gastrointestinal, and genito-urinary tract, as well as adenocarcinomas which include malignancies such as most colon cancers, renal-cell carcinoma, prostate cancer and/or testicular tumors, non-small cell carcinoma of the lung, cancer of the small intestine and cancer of the esophagus.

The term "carcinoma" is art recognized and refers to malignancies of epithelial or endocrine tissues including respiratory system carcinomas, gastrointestinal system carcinomas, genitourinary system carcinomas, testicular carcinomas, breast carcinomas, prostatic carcinomas, endocrine system carcinomas, and melanomas. Exemplary carcinomas include those forming from tissue of the cervix, lung, prostate, breast, head and neck, colon and ovary. The term also includes carcinosarcomas, e.g., which include malignant tumors composed of carcinomatous and sarcomatous tissues. An "adenocarcinoma" refers to a carcinoma derived from glandular tissue or in which the tumor cells form recognizable glandular structures.

The term "sarcoma" is art recognized and refers to malignant tumors of mesenchymal derivation.

Additional examples of proliferative disorders include hematopoietic neoplastic disorders. As used herein, the term "hematopoietic neoplastic disorders"

includes diseases involving hyperplastic/neoplastic cells of hematopoietic origin, e.g., arising from myeloid, lymphoid or erythroid lineages, or precursor cells thereof.

Preferably, the diseases arise from poorly differentiated acute leukemias, e.g., erythroblastic leukemia and acute megakaryoblastic leukemia. Additional exemplary  
5 myeloid disorders include, but are not limited to, acute promyeloid leukemia (APML), acute myelogenous leukemia (AML) and chronic myelogenous leukemia (CML) (reviewed in Vaickus, L. (1991) *Crit Rev. in Oncol./Hematol.* 11:267-97); lymphoid malignancies include, but are not limited to acute lymphoblastic leukemia (ALL) which includes B-lineage ALL and T-lineage ALL, chronic lymphocytic leukemia  
10 (CLL), prolymphocytic leukemia (PLL), hairy cell leukemia (HLL) and Waldenstrom's macroglobulinemia (WM). Additional forms of malignant lymphomas include, but are not limited to non-Hodgkin lymphoma and variants thereof, peripheral T cell lymphomas, adult T cell leukemia/lymphoma (ATL), cutaneous T-cell lymphoma (CTCL), large granular lymphocytic leukemia (LGL), Hodgkin's  
15 disease and Reed-Sternberg disease.

In general, the engineered RNA precursors of the invention are designed to target genes associated with particular disorders. Examples of such genes associated with proliferative disorders that can be targeted include activated ras, p53, BRCA-1, and BRCA-2. Other specific genes that can be targeted are those associated with  
20 amyotrophic lateral sclerosis (ALS; e.g., superoxide dismutase-1 (SOD1)); Huntington's disease (e.g., huntingtin), Parkinson's disease (parkin), and genes associated with autosomal dominant disorders.

The engineered RNAs of the invention can be used to treat a variety of immune disorders, in particular those associated with overexpression of a gene or  
25 expression of a mutant gene. Examples of hematopoietic disorders or diseases include, but are not limited to, autoimmune diseases (including, for example, diabetes mellitus, arthritis (including rheumatoid arthritis, juvenile rheumatoid arthritis, osteoarthritis, psoriatic arthritis), multiple sclerosis, encephalomyelitis, myasthenia gravis, systemic lupus erythematosus, autoimmune thyroiditis, dermatitis (including  
30 atopic dermatitis and eczematous dermatitis), psoriasis, Sjögren's Syndrome, Crohn's disease, aphthous ulcer, iritis, conjunctivitis, keratoconjunctivitis, ulcerative colitis, asthma, allergic asthma, cutaneous lupus erythematosus, scleroderma, vaginitis,

proctitis, drug eruptions, leprosy reversal reactions, erythema nodosum leprosum, autoimmune uveitis, allergic encephalomyelitis, acute necrotizing hemorrhagic encephalopathy, idiopathic bilateral progressive sensorineural hearing loss, aplastic anemia, pure red cell anemia, idiopathic thrombocytopenia, polychondritis, Wegener's  
5 granulomatosis, chronic active hepatitis, Stevens-Johnson syndrome, idiopathic sprue, lichen planus, Graves' disease, sarcoidosis, primary biliary cirrhosis, uveitis posterior, and interstitial lung fibrosis), graft-versus-host disease, cases of transplantation, and allergy such as, atopic allergy.

Examples of disorders involving the heart or "cardiovascular disorder"

10 include, but are not limited to, a disease, disorder, or state involving the cardiovascular system, *e.g.*, the heart, the blood vessels, and/or the blood. A cardiovascular disorder can be caused by an imbalance in arterial pressure, a malfunction of the heart, or an occlusion of a blood vessel, *e.g.*, by a thrombus. Examples of such disorders include hypertension, atherosclerosis, coronary artery  
15 spasm, congestive heart failure. coronary artery disease, valvular disease, arrhythmias, and cardiomyopathies.

Disorders which may be treated by methods described herein include, but are not limited to, disorders associated with an accumulation in the liver of fibrous tissue, such as that resulting from an imbalance between production and degradation of the  
20 extracellular matrix accompanied by the collapse and condensation of preexisting fibers.

Additionally, molecules of the invention can be used to treat viral diseases, including but not limited to hepatitis B, hepatitis C, herpes simplex virus (HSV), HIV-AIDS, poliovirus, and smallpox virus. Molecules of the invention are engineered as  
25 described herein to target expressed sequences of a virus, thus ameliorating viral activity and replication. The molecules can be used in the treatment and/or diagnosis of viral infected tissue. Also, such molecules can be used in the treatment of virus-associated carcinoma, such as hepatocellular cancer.

### 30 Uses of Engineered RNA Precursors to Induce RNAi

Engineered RNA precursors, introduced into cells or whole organisms as described herein, will lead to the production of a desired siRNA molecule. Such an

siRNA molecule will then associate with endogenous protein components of the RNAi pathway to bind to and target a specific mRNA sequence for cleavage and destruction. In this fashion, the mRNA to be targeted by the siRNA generated from the engineered RNA precursor will be depleted from the cell or organism, leading to a decrease in the concentration of the protein encoded by that mRNA in the cell or organism.

For example, one may be seeking to discover a small molecule that reduces the activity of a kinase whose overexpression leads to unrestrained cell proliferation. This kinase is overexpressed in a variety of cancer cells. A key question to be determined is whether or not decreasing the activity of this kinase in adult mammals would have unexpected deleterious effects. By expressing an engineered RNA precursor that targets for destruction by the RNAi pathway the mRNA encoding the kinase throughout the tissues of an adult mouse, the deleterious effects of such a potential drug can be determined. That is, the method described here will allow rapid assessment of the suitability of the kinase as a drug target.

The new nucleic acid molecules that encode the engineered RNA precursors can also be used to create large numbers of cells or vectors in microarrays in which each cell or vector in the array includes nucleic acid molecules that encode an engineered RNA precursor that is specific for a different target gene. See, e.g., Ziauddin et al., *Nature*, 411:107-110 (2001).

### EXAMPLES

The invention is further described in the following examples, which do not limit the scope of the invention described in the claims.

Example 1 – Producing an Engineered RNA Precursor

To produce an engineered RNA precursor that will target a gene for firefly luciferase for cleavage, the sequence of the coding portion of the mRNA from firefly luciferase was examined to select a suitable sequence. At a position more than 100  
 5 nucleotides, but less than 300 nucleotides, 3' to the start of translation, the sequence CGUACGCGGAAUACU-UCGAUU (SEQ ID NO:16) was found immediately after the sequence AA. Since this sequence meets the criteria for selection of an siRNA, it was chosen. An engineered precursor RNA was then designed that includes this sequence (underlined) as represented below (and shown in FIG. 2B):

10

5'-

GGCAAACCGUACGCGGAAUACUUCGAUUAGUAAUUACACAUCAUAA-  
 UCGAAG UAUUCCGCGUACGUUUGCU-3' (SEQ ID NO:2)

15

Synthetic deoxyoligonucleotide sequences were prepared that serve as an *in vitro* transcription template for preparing this engineered RNA precursor using the enzyme T7 RNA Polymerase according to published protocols. Below are shown the two deoxyoligonucleotides. The oligonucleotides contain the sequence of the T7 RNA Polymerase promoter (underlined in the top strand) to facilitate *in vitro*  
 20 transcription into RNA.

Top oligo:

5'-

GCGTAATACGACTCACTATAGGCAAACGTACGCGGAATACTTCGAT-

25

TAGTAATTACACATCATAATCGAAGTATTCGCGTACGTTTGCT-3' (SEQ ID  
 NO:17)

Bottom oligo:

5'-

TGTAGTCACGTACGCGGAATACTTCGAAGAAACGAGTAATTACTAAAT-

30

CGAA GTATTCGCGTACGTTTGCCTATAGTGAGTCGTATTACGC-3' (SEQ ID  
 NO:18)

Next, the double-stranded DNA was formed by annealing the two deoxyoligonucleotides and was transcribed into RNA using T7 RNA Polymerase. The resulting engineered RNA precursor was purified by standard means, and then tested for its ability to promote cleavage of the target mRNA *in vitro* in a standard RNAi reaction.

Briefly, firefly luciferase mRNA was prepared by *in vitro* transcription and radiolabeled using  $\alpha$ -<sup>32</sup>P-GTP and Guanylyl Transferase as described previously (Tuschl *et al.*, (1999) cited *supra*) to generate a radioactive target mRNA. The radioactive target mRNA was incubated in a standard *in vitro* RNAi reaction with *Drosophila* embryo lysate and 50 nM engineered RNA precursor (ESP) for 0 and 3 hours as described previously (Tuschl *et al.*, (1999) cited *supra*). The reaction products were isolated and analyzed by denaturing acrylamide gel electrophoresis as described previously (Zamore *et al.*, (2000) cited *supra*). As shown in FIG 3, the engineered RNA precursor induced sequence-specific cleavage of the radioactive target mRNA (5' cleavage product). Thus, the precursor was shown to mediate RNAi. FIG 3 also shows the RNA cleavage product of a standard siRNA (also incubated for 0 and 3 hours), which produced the same sequence-specific 5' cleavage product as the ESP.

#### Example 2 - Engineered *let-7* Precursor RNAs Asymmetrically Trigger RNAi *in vitro*

Two engineered *let-7* RNA precursors (ESPs) were prepared in which the stem of pre-*let-7* (FIGs. 4B and 4C) was altered to contain a sequence complementary to firefly luciferase. Because most stRNAs begin with uracil, the ESPs were designed so that the luciferase-complementary sequence (anti-sense luciferase) began with U. Since stRNAs can be encoded on either the 5' or the 3' side of the precursor hairpin (e.g., on either stem) the anti-sense luciferase sequence (in bold) was placed on the 3' side of the stem in one ESP (3' ESP) (FIG. 4B) and on the 5' side in the second stem (5' ESP) (FIG. 4C).

The ESP RNAs were prepared as generally described in Example 1 by using the following DNA oligonucleotide pairs to generate partially single-stranded T7 RNA Polymerase transcription templates:

5'-GTAATACGACTCACTATAG-3' (SEQ ID NO:19)

5'-

GGCAAATTCGAAGTATTCCGCGTACGTGATGATGTGTAATTACTCACG  
5 TACGCGGAATACTTCGAATTTGCCTATAGTGAGTCGTATTAC-3' (5' ESP)  
(SEQ ID NO:20)

5'-

GGCAAATCGTACGCGGAATACTTCGAAAATGATGTGTAATTACTTTTC  
10 GAAGTATTCCGCGTACGATTTCCTATAGTGAGTCGTATTAC-3' (3' ESP)  
(SEQ ID NO:21)

The anti-sense firefly luciferase target RNA has been described previously (A. Nykänen, B. Haley, P. D. Zamore, Cell 107, 309, 2001).

15 The ability of each ESP to direct luciferase-specific RNAi in an *in vitro* reaction was tested against a target mRNA (shown schematically in FIG. 4D) containing a portion of the firefly luciferase mRNA and a sequence fully complementary to *let-7* (the target was constructed by standard techniques and synthesized using T7 RNA Polymerase). As a control, an siRNA duplex containing  
20 the anti-sense luciferase sequence was used (FIG. 4A).

FIG. 4E is an autoradiograph showing the results of an assay for determining whether the 5' and 3' ESPs can equally promote cleavage of the target RNA *in vitro* (the assay conditions are described in Example 1, except the ESPs and control were incubated for 0 and 2 hours). Both the 3' and the 5' ESPs directed cleavage of the  
25 target RNA within the luciferase sequences, the same site cleaved when the RNAi reaction was programmed with the control siRNA.

### Example 3 – Preparing a Transgene Encoding an Engineered RNA Precursor

To prepare a transgene encoding an engineered precursor to target destruction  
30 of the luciferase mRNA in a transgenic mouse that expresses firefly luciferase mRNA in all of its cells, the engineered RNA precursor sequence described in Example 1 is cloned by standard recombinant DNA methods into a nucleic acid molecule, e.g., a

vector containing a constitutively expressed promoter sequence and the desired nucleic acid sequence (transgene) encoding the engineered RNA precursor as illustrated in FIG 5. This vector will also contain sequences appropriate for its introduction into ES cells to produce transgenic mice by standard methods. The resulting transgene expresses the engineered RNA precursor in all cells of a transgenic mouse.

The engineered precursor RNA is then processed by Dicer and other components of the RNAi machinery to yield an siRNA directed against the firefly luciferase gene. This siRNA directs cleavage of the luciferase mRNA, resulting in a decrease in the expression of luciferase mRNA in the cells of the animal.

The same methods can be used to silence other target genes, either using constitutively expressed or inducible expression systems in a variety of transgenic animals.

## OTHER EMBODIMENTS

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.



**WHAT IS CLAIMED IS:**

1. An isolated nucleic acid molecule comprising
  - 5 a regulatory sequence operably linked to a nucleic acid sequence that encodes an engineered ribonucleic acid (RNA) precursor, wherein the precursor comprises
    - (i) a first stem portion comprising a sequence of at least 18 nucleotides that is complementary to a sequence of a messenger RNA (mRNA) of a target  
10 gene;
    - (ii) a second stem portion comprising a sequence of at least 18 nucleotides that is sufficiently complementary to the first stem portion to hybridize with the first stem portion to form a duplex stem; and
    - (iii) a loop portion that connects the two stem portions.
- 15 2. The nucleic acid molecule of claim 1, wherein the first stem portion is fully complementary to the mRNA sequence.
3. The nucleic acid molecule of claim 1, wherein the second stem portion is fully  
20 complementary to the first stem portion.
4. The nucleic acid molecule of claim 1, wherein the first stem portion is located at a 5' end of the RNA precursor.
- 25 5. The nucleic acid molecule of claim 1, wherein the first stem portion is located at a 3' end of the RNA precursor.
6. The nucleic acid molecule of claim 1, wherein the loop portion comprises at least 4 nucleotides.
- 30 7. The nucleic acid molecule of claim 1, wherein the loop portion comprises at least 7 nucleotides.

8. The nucleic acid molecule of claim 1, wherein the loop portion comprises 11 nucleotides.
- 5 9. The nucleic acid molecule of claim 1, wherein the sequence of the mRNA is located from 100 to 300 nucleotides 3' of the start of translation of the mRNA.
10. The nucleic acid molecule of claim 1, wherein the sequence of the mRNA is located in a 5' untranslated region (UTR) or a 3' UTR of the mRNA.
- 10 11. The nucleic acid molecule of claim 1, wherein the first and second stem portions each comprise about 18 to about 30 nucleotides.
12. The nucleic acid molecule of claim 1, wherein the first and second stem portions each comprise about 22 to about 28 nucleotides.
- 15 13. The nucleic acid molecule of claim 1, wherein the first and second stem portions each comprise the same number of nucleotides.
14. The nucleic acid molecule of claim 1, wherein one of the first and second stem portions comprises 1 to 4 more nucleotides than the other stem portion.
- 20 15. The nucleic acid molecule of claim 1, wherein the regulatory sequence comprises a Pol III or Pol II promoter.
- 25 16. The nucleic acid molecule of claim 1, wherein the regulatory sequence is constitutive or inducible.
17. The nucleic acid molecule of claim 1, wherein the sequence of the engineered RNA precursor is the sequence set forth in SEQ ID NO:1, 2, 3, 4, 5, 8, or 9.
- 30

18. The nucleic acid molecule of claim 1, comprising the sequence set forth in SEQ ID NO:10, 11, 17, 18, 20, or 21, or a complement thereof.
19. A vector comprising the nucleic acid molecule of claim 1.
20. The vector of claim 19, wherein the vector is a plasmid or a viral vector.
21. The vector of claim 20, wherein the viral vector is a retroviral vector.
22. A host cell containing the nucleic acid molecule of claim 1.
23. The host cell of claim 22, wherein the cell is a mammalian cell.
24. A transgene comprising the nucleic acid of claim 1.
25. An engineered RNA precursor comprising
- (i) a first stem portion comprising a sequence of at least 18 nucleotides that is complementary to a sequence of a messenger RNA (mRNA) of a target gene;
  - (ii) a second stem portion comprising a sequence of at least 18 nucleotides that is sufficiently complementary to the first stem portion to hybridize with the first stem portion to form a duplex stem; and
  - (iii) a loop portion that connects the two stem portions.
26. The precursor of claim 25, wherein the first stem portion is fully complementary to the mRNA sequence.
27. The precursor of claim 25, wherein the second stem portion is fully complementary to the first stem portion.
28. The precursor of claim 25, wherein the first stem portion is located at a 5' end of the RNA precursor.

29. The precursor of claim 25, wherein the first stem portion is located at a 3' end of the RNA precursor.
- 5 30. The precursor of claim 25, wherein the loop portion comprises at least 4 nucleotides.
31. The precursor of claim 25, wherein the loop portion comprises at least 7 nucleotides.
- 10 32. The precursor of claim 25, wherein the loop portion comprises 11 nucleotides.
33. The precursor of claim 25, wherein the sequence of the mRNA is located in a 5' untranslated region (UTR) or a 3' UTR of the mRNA.
- 15 34. The precursor of claim 25, wherein the first and second stem portions each comprise about 18 to about 30 nucleotides.
35. The precursor of claim 25, wherein the first and second stem portions each comprise about 22 to about 28 nucleotides.
- 20 36. The precursor of claim 25, wherein the first and second stem portions each comprise the same number of nucleotides.
37. The precursor of claim 25, wherein one of the first and second stem portions comprises 1 to 4 more nucleotides than the other stem portion.
- 25 38. The precursor of claim 25, comprising the sequence set forth in SEQ ID NO:1, 2, 3, 4, 5, 8, or 9.
- 30 39. The precursor of claim 25, wherein the target gene is a human gene.
40. The precursor of claim 25, wherein the target gene is a mutant human gene.

41. The precursor of claim 25, wherein the target gene is a viral gene.
- 5 42. A transgenic, non-human animal, one or more of whose cells comprise a transgene comprising a nucleic acid molecule of claim 1, wherein the transgene is expressed in one or more cells of the transgenic animal resulting in the animal exhibiting ribonucleic acid interference (RNAi) of the target gene by the engineered RNA precursor.
- 10 43. The transgenic animal of claim 42, wherein the transgene is expressed in one or more cardiac cells, lymphocytes, liver cells, vascular endothelial cells, or spleen cells.
- 15 44. The transgenic animal of claim 42, wherein the regulatory sequence is constitutive or inducible.
45. The transgenic animal of claim 42, wherein the regulatory sequence is tissue specific.
- 20 46. The transgenic animal of claim 42, wherein the regulatory sequence is a Pol III or Pol II promoter.
47. The transgenic animal of claim 42, wherein the regulatory sequence is an exogenous sequence.
- 25 48. The transgenic animal of claim 42, wherein the animal is a non-human primate.
49. The transgenic animal of claim 42, wherein the animal is a rodent.
- 30 50. The transgenic animal of claim 42, wherein the animal is a mouse.

51. A cell derived from the transgenic animal of claim 42.
52. The cell of claim 51, wherein the cell is a lymphocyte, a liver cell, a cardiac cell, a vascular endothelial cell, or a spleen cell.
53. A method of inducing ribonucleic acid interference (RNAi) of a target gene in a cell in an animal, the method comprising  
obtaining a transgenic animal of claim 42 comprising a transgene comprising a nucleic acid molecule encoding an engineered RNA precursor and an inducible promoter; and  
inducing the cell to express the precursor to form a small interfering ribonucleic acid (siRNA) within the cell, thereby inducing RNAi of the target gene in the animal.
54. A method of inducing ribonucleic acid interference (RNAi) of a target gene in a cell, the method comprising  
obtaining a host cell of claim 22;  
culturing the cell; and  
enabling the cell to express the RNA precursor to form a small interfering ribonucleic acid (siRNA) within the cell, thereby inducing RNAi of the target gene in the cell.

FIG. 1

*stRNA pathway*

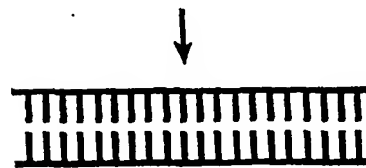
developmentally  
regulated transcription



precursor

*RNAi pathway*

experimentally  
introduced dsRNA



dsRNA

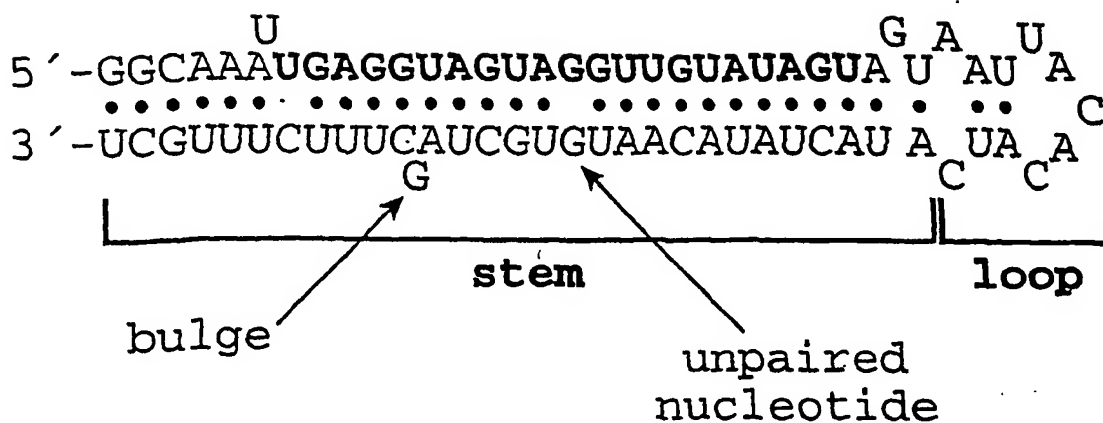
**Dicer**ATP → ADP + P<sub>i</sub>

stRNA



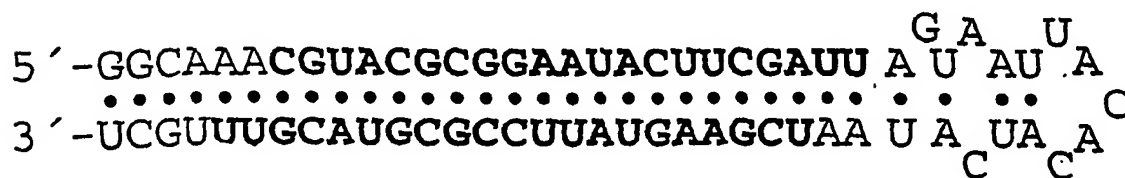
siRNA

FIG. 2A



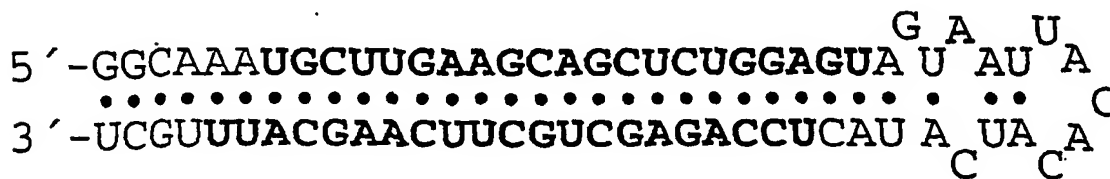
(SEQ ID NO:1)

FIG. 2B



(SEQ ID NO:2)

FIG. 2C



(SEQ ID NO:3)



FIG. 2D

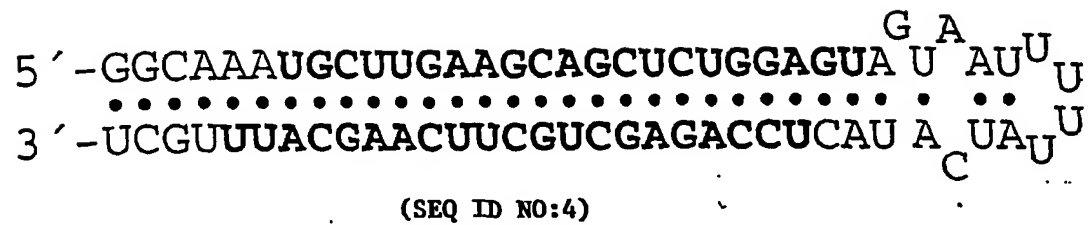
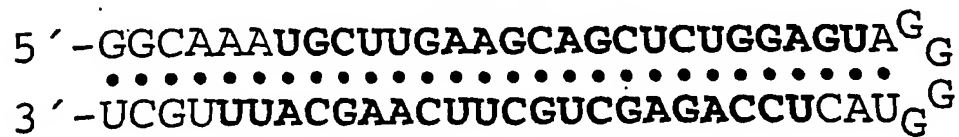


FIG. 2E



(SEQ ID NO:5)

FIG. 3

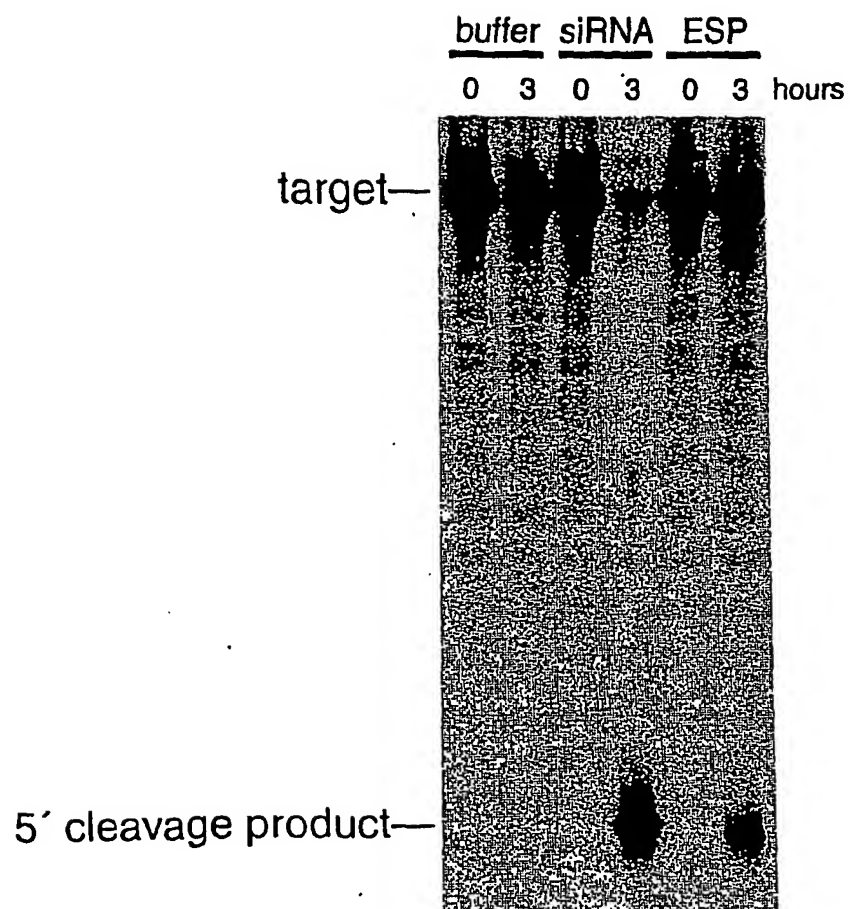
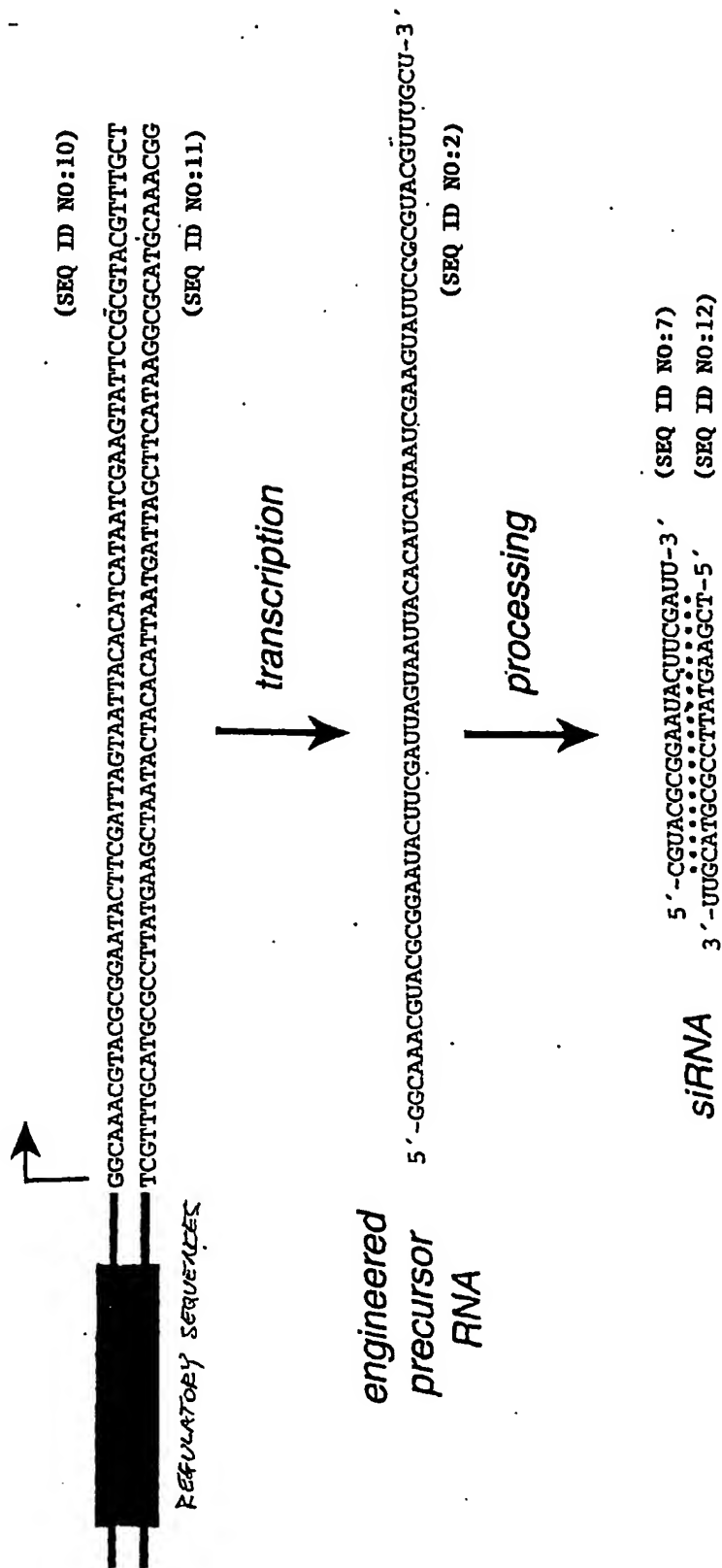




FIG. 5



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/22010

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C07H 21/02, 21/04; C12N 5/00; A01K 67/00

US CL : 536/23.1, 24.5; 435/325; 800/8

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1, 24.5; 435/325; 800/8

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched reviewed 60/305,185

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
STN, EAST

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 6,022,863 A (PEYMAN) 08 February 2000 (08.02.2000), entire reference, in particular examples in figures	1-16, 19-37, 39-54
X	US 5,631,146 A (SZOSTAK et al.) 20 May 1997 (20.05.1997), entire reference.	1-16, 19-37, 39-54
X,E	US 2001/0008771 A1 (SEIBEL et al.) 19 July 2001 (19.07.2001), entire reference.	1-16, 19-37, 39-54

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"B" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;"

document member of the same patent family

Date of the actual completion of the international search

13 November 2002 (13.11.2002)

Date of mailing of the international search report

04 DEC 2002

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. 703 305-3230

Authorized officer

Selicia D. Roberts Jfor  
Joseph Wolfach

Telephone No. 703 308-1235

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/22010

## Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claim Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claim Nos.: 17,18 and 38  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
The claims are drawn to specific SEQ ID NOs which have not been provided.
3. ☐ Claim Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of Item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
  2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
  3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
- Remark on Protest ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.